

## **RAW SEQUENCE LISTING**

— The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/534,171  
Source: JFWP  
Date Processed by STIC: 08/04/2006

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 08/04/2006

PATENT APPLICATION: US/10/534,171

TIME: 12:55:59

Input Set : E:\3912.1000-000 seq list.txt

Output Set: N:\CRF4\08042006\J534171.raw

```

4 <110> APPLICANT: Porro, Danilo
5      Branduardi, Paola
6      Valli, Minoska
7      Alberghina, Lilia
9 <120> TITLE OF INVENTION: Process for Expression and Secretion of
10     Proteins by the Non-Conventional Yeast Zygosaccharomyces
11     Bailii
13 <130> FILE REFERENCE: 3912.1000-000
15 <140> CURRENT APPLICATION NUMBER: 10/534,171
16 <141> CURRENT FILING DATE: 2005-05-06
18 <150> PRIOR APPLICATION NUMBER: PCT/EP2003/012377
19 <151> PRIOR FILING DATE: 2003-11-06
21 <150> PRIOR APPLICATION NUMBER: 102 52 245.6
22 <151> PRIOR FILING DATE: 2002-11-07
24 <160> NUMBER OF SEQ ID NOS: 95
26 <170> SOFTWARE: FastSEQ for Windows Version 4.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 57
30 <212> TYPE: DNA
31 <213> ORGANISM: Saccharomyces cerevisiae
33 <400> SEQUENCE: 1
34 cgacagtataa ttttgccgaa tttaatatgct tctactgaaa aacagtggac catgtga      57
36 <210> SEQ ID NO: 2
37 <211> LENGTH: 19
38 <212> TYPE: PRT
39 <213> ORGANISM: Saccharomyces cerevisiae
41 <400> SEQUENCE: 2
42 Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
43 1          5          10          15
44 Ala Leu Ala
48 <210> SEQ ID NO: 3
49 <211> LENGTH: 255
50 <212> TYPE: DNA
51 <213> ORGANISM: Saccharomyces cerevisiae
53 <400> SEQUENCE: 3
54 cgacagtataa ttttgccgaa tttaatatgct tctactgaaa aacagtggac catgtgaaaa 60
55 gatgcatctc atttatcaaa cacataatat tcaagtgagc cttacttcaa ttgtattgaa 120
56 gtgcaagaaa accaaaaagc aacaacaggt tttggataag tacatatata agagggcctt 180
57 ttgttcccat caaaaatgtt actgttctta cgattcatTT acgattcaag aatagttcaa 240
58 acaagaagat tacaa                                     255
60 <210> SEQ ID NO: 4
61 <211> LENGTH: 85
62 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/534,171

DATE: 08/04/2006

TIME: 12:55:59

Input Set : E:\3912.1000-000 seq list.txt

Output Set: N:\CRF4\08042006\J534171.raw

```

63 <213> ORGANISM: Saccharomyces cerevisiae
64 <400> SEQUENCE: 4
65 Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
66 1 5 10 15
67 Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
68 20 25 30
69 Ile Pro Ala Glu Ala Val Ile Gly Tyr Leu Asp Leu Glu Gly Asp Phe
70 35 40 45
71 Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
72 50 55 60
73 Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
74 65 70 75 80
75 Ser Leu Asp Lys Arg
76 85
80 <210> SEQ ID NO: 5
81 <211> LENGTH: 63
82 <212> TYPE: DNA
83 <213> ORGANISM: Aspergillus niger
84 <400> SEQUENCE: 5
85 atgatggtcg cgtggtggtc tctatttctg tacggccttc aggtcgcggc acctgctttg 60
86 gct 63
89 <210> SEQ ID NO: 6
90 <211> LENGTH: 21
91 <212> TYPE: PRT
92 <213> ORGANISM: Aspergillus niger
93 <400> SEQUENCE: 6
94 Met Met Val Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
95 1 5 10 15
96 Ala Pro Ala Leu Ala
97 20
101 <210> SEQ ID NO: 7
102 <211> LENGTH: 93
103 <212> TYPE: DNA
104 <213> ORGANISM: Bacillus sp.
105 <400> SEQUENCE: 7
106 atgcaaaaca cagcgaaaaa ctccatctgg cagagggtgc gccacagcgc cattgcctta 60
107 tccgctctca gtttaccctt tggcctgcag gcc 93
110 <210> SEQ ID NO: 8
111 <211> LENGTH: 31
112 <212> TYPE: PRT
113 <213> ORGANISM: Bacillus sp.
114 <400> SEQUENCE: 8
115 Met Gln Asn Thr Ala Lys Asn Ser Ile Trp Gln Arg Val Arg His Ser
116 1 5 10 15
117 Ala Ile Ala Leu Ser Ala Leu Ser Leu Ser Phe Gly Leu Gln Ala
118 20 25 30
122 <210> SEQ ID NO: 9
123 <211> LENGTH: 63
124 <212> TYPE: DNA

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/534,171

DATE: 08/04/2006

TIME: 12:55:59

Input Set : E:\3912.1000-000 seq list.txt

Output Set: N:\CRF4\08042006\J534171.raw

```

125 <213> ORGANISM: Aspergillus oryzae
127 <400> SEQUENCE: 9
128 atgcagtcca tcaagcgtag cttgctcttc ctggagcta tccttccgcg ggtcctcggt 60
129 gcc 63
131 <210> SEQ ID NO: 10
132 <211> LENGTH: 21
133 <212> TYPE: PRT
134 <213> ORGANISM: Aspergillus oryzae
136 <400> SEQUENCE: 10
137 Met Gln Ser Ile Lys Arg Thr Leu Leu Leu Gly Ala Ile Leu Pro
138 1 5 10 15
139 Ala Val Leu Gly Ala
140 20
143 <210> SEQ ID NO: 11
144 <211> LENGTH: 75
145 <212> TYPE: DNA
146 <213> ORGANISM: Bacillus amyloliquefaciens
148 <400> SEQUENCE: 11
149 atgaaacgag tggtgcta tcttgctacc ggattgttta tgagtttggtg tgggatcact 60
150 tctagtgttt cggct 75
152 <210> SEQ ID NO: 12
153 <211> LENGTH: 25
154 <212> TYPE: PRT
155 <213> ORGANISM: Bacillus amyloliquefaciens
157 <400> SEQUENCE: 12
158 Met Lys Arg Val Leu Leu Ile Leu Val Thr Gly Leu Phe Met Ser Leu
159 1 5 10 15
160 Cys Gly Ile Thr Ser Ser Val Ser Ala
161 20 25
164 <210> SEQ ID NO: 13
165 <211> LENGTH: 51
166 <212> TYPE: DNA
167 <213> ORGANISM: Saccharomycopsis fibuligera
169 <400> SEQUENCE: 13
170 atgttgatga tagtacagct tttggtcttt gcactaggcc ttgctgttgc t 51
172 <210> SEQ ID NO: 14
173 <211> LENGTH: 17
174 <212> TYPE: PRT
175 <213> ORGANISM: Saccharomycopsis fibuligera
177 <400> SEQUENCE: 14
178 Met Leu Met Ile Val Gln Leu Leu Val Phe Ala Leu Gly Leu Ala Val
179 1 5 10 15
180 Ala
184 <210> SEQ ID NO: 15
185 <211> LENGTH: 51
186 <212> TYPE: DNA
187 <213> ORGANISM: Saccharomycopsis fibuligera
189 <400> SEQUENCE: 15
190 atgttgatga ttttgaact cttagtactt attatagggc ttggagttgc t 51

```

## RAW SEQUENCE LISTING

DATE: 08/04/2006

PATENT APPLICATION: US/10/534,171

TIME: 12:55:59

Input Set : E:\3912.1000-000 seq list.txt

Output Set: N:\CRF4\08042006\J534171.raw

```

192 <210> SEQ ID NO: 16
193 <211> LENGTH: 17
194 <212> TYPE: PRT
195 <213> ORGANISM: Saccharomycopsis fibuligera
197 <400> SEQUENCE: 16
198 Met Leu Leu Ile Leu Glu Leu Leu Val Leu Ile Ile Gly Leu Gly Val
199 1 5 10 15
200 Ala
204 <210> SEQ ID NO: 17
205 <211> LENGTH: 66
206 <212> TYPE: DNA
207 <213> ORGANISM: Hypocrea pecorina
209 <400> SEQUENCE: 17
210 atggcgccct cagttacact gccgttgacc acggccatcc tggccattgc cgggctcgtc 60
211 gccgcc 66
213 <210> SEQ ID NO: 18
214 <211> LENGTH: 22
215 <212> TYPE: PRT
216 <213> ORGANISM: Hypocrea pecorina
218 <400> SEQUENCE: 18
219 Met Ala Pro Ser Val Thr Leu Pro Leu Thr Thr Ala Ile Leu Ala Ile
220 1 5 10 15
221 Ala Arg Leu Val Ala Ala
222 20
225 <210> SEQ ID NO: 19
226 <211> LENGTH: 63
227 <212> TYPE: DNA
228 <213> ORGANISM: Hypocrea pecorina
230 <400> SEQUENCE: 19
231 atgaacaagt ccgtggctcc attgctgctt gcagcggtcca tactatatgg cggcgccgctc 60
232 gca 63
234 <210> SEQ ID NO: 20
235 <211> LENGTH: 21
236 <212> TYPE: PRT
237 <213> ORGANISM: Hypocrea pecorina
239 <400> SEQUENCE: 20
240 Met Asn Lys Ser Val Ala Pro Leu Leu Leu Ala Ala Ser Ile Leu Tyr
241 1 5 10 15
242 Gly Gly Ala Val Ala
243 20
246 <210> SEQ ID NO: 21
247 <211> LENGTH: 48
248 <212> TYPE: DNA
249 <213> ORGANISM: Arxula adeninivorans
251 <400> SEQUENCE: 21
252 atgcgtcagt ttctagcact tgctgctgct gcttccatag cgggtggca 48
254 <210> SEQ ID NO: 22
255 <211> LENGTH: 18
256 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/534,171

DATE: 08/04/2006

TIME: 12:55:59

Input Set : E:\3912.1000-000 seq list.txt

Output Set: N:\CRF4\08042006\J534171.raw

257 <213> ORGANISM: Arxula adeninivorans  
 259 <400> SEQUENCE: 22  
 260 Met Arg Gln Phe Leu Ala Leu Ala Ala Ala Ser Ile Ala Val Ala  
 261 1 5 10 15  
 262 Asp Ser  
 266 <210> SEQ ID NO: 23  
 267 <211> LENGTH: 57  
 268 <212> TYPE: DNA  
 269 <213> ORGANISM: Homo sapiens  
 271 <400> SEQUENCE: 23  
 272 atgcagcgac tatgtgtgta tgtgctgac tttgcactgg ctctggccgc cttctct 57  
 274 <210> SEQ ID NO: 24  
 275 <211> LENGTH: 19  
 276 <212> TYPE: PRT  
 277 <213> ORGANISM: Homo sapiens  
 279 <400> SEQUENCE: 24  
 280 Met Gln Arg Leu Cys Val Tyr Val Leu Ile Phe Ala Leu Ala Leu Ala  
 281 1 5 10 15  
 282 Ala Phe Ser  
 286 <210> SEQ ID NO: 25  
 287 <211> LENGTH: 75  
 288 <212> TYPE: DNA  
 289 <213> ORGANISM: Rhizopus oryzae  
 291 <400> SEQUENCE: 25  
 292 atgcaactgt tcaatttgcc attgaaagtt tcattctttc tcgtcctctc ttacttttct 60  
 293 ttgctcgttt ctgct 75  
 295 <210> SEQ ID NO: 26  
 296 <211> LENGTH: 25  
 297 <212> TYPE: PRT  
 298 <213> ORGANISM: Rhizopus oryzae  
 300 <400> SEQUENCE: 26  
 301 Met Gln Leu Phe Asn Leu Pro Leu Lys Val Ser Phe Phe Leu Val Leu  
 302 1 5 10 15  
 303 Ser Tyr Phe Ser Leu Leu Val Ser Ala  
 304 20 25  
 307 <210> SEQ ID NO: 27  
 308 <211> LENGTH: 48  
 309 <212> TYPE: DNA  
 310 <213> ORGANISM: Aspergillus niger  
 312 <400> SEQUENCE: 27  
 313 atgcagactc tccttgtag ctgcttgtag gtctccctcg ctgcggcc 48  
 315 <210> SEQ ID NO: 28  
 316 <211> LENGTH: 16  
 317 <212> TYPE: PRT  
 318 <213> ORGANISM: Aspergillus niger  
 320 <400> SEQUENCE: 28  
 321 Met Gln Thr Leu Leu Val Ser Ser Leu Val Val Ser Leu Ala Ala Ala  
 322 1 5 10 15  
 325 <210> SEQ ID NO: 29

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/534,171

DATE: 08/04/2006

TIME: 12:56:00

Input Set : E:\3912.1000-000 seq list.txt

Output Set: N:\CRF4\08042006\J534171.raw

L:1286 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:78

L:1309 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:79